UNIVERSITETET I OSLO

Matematisk Institutt

EXAM IN:	**STK 4900/9900 –** **Statistical Methods and Applications**
WITH:	**Nils Lid Hjort**
AUXILIA:	**All printed material, also the candidate's own notes** **Calculator**
TIME FOR EXAM:	**Friday 7/vi/2024, 15:00–19:00**

Notes for the solutions, as per 7/vi/2024

Exercise 1: happy birthday

(a) The situation fits the binomial setup, provided times for deaths are seen as independent in the population. The p is seen as the probability of dying in the time period of three months before one's birthday date. Disregarding the small factor that the months do not have precisely the same number of days, we should have $p = 3/12 = 0.25$ if people's death days are independent of their dates of birth.

(b) The X is approximately normal, and with $p = 0.25$ the mean is $\xi = np = 186.75$ and the standard deviation $\sigma = (np(1-p))^{1/2} = 11.834$. So about 95 percent of outcomes will be inside $\xi \pm 1.96\,\sigma$, i.e. inside $[163.5, 209.9]$, or $[164, 210]$ rounded to numbers.

(c) We should reject the hypothesis $p = 0.25$, since the outcome $X = 60$ is very clearly too low. So p is significantly lower, which means a significant part of the population holds on to their lives enough to avoid dying in this pre-birthday period.

Exercise 2: adiposity, LDL, and heart disease

(a) Note that the setup, and the questions being asked, are rather similar to one of the Oblig II exercises. The first question is a matter of computing the ratios $z =$ estimate/sterror, for the parameters $\beta_0, \ldots, \beta_4$, and then checking which of these might be too big in absolute value, i.e. outside the range of the standard normal. It is seen that $z_0 = -3.9384/0.4761$ is big enough in absolute value, for the intercept, and likewise $z_1 = 0.1266/0.0190 = 6.6632$ for β_1; the other ratios are rather small and within normal range. That β_1 is significantly positive is also clear from the figure.

(b) We plug in values for Mr. Jones, $x = (1, 30, 50, 20, 25)$, leading to estimated value of $\gamma_J = \beta_0 + \beta_1 x_1 + \cdots + \beta_4 x_4$ equal to $\widehat{\gamma}_J = 0.3331$, and then to $\widehat{p}_J = \exp(0.3331)/(1 + \exp(0.3331)) = 0.582$.

(c) The confidence recipe, used in Oblig II in a similar problem, is to use the normal approximation $\mathrm{N}_5(\beta, \Sigma)$ for the vector of estimators $\widehat{\beta}$, for a 5×5 matrix Σ that we get from running `glm` with logistic regression. This leads to $\widehat{\gamma}_J$ above being approximately normal, with a variance $\widehat{\tau}_J^2$ we can compute, and then the usual confidence interval

$\widehat\gamma_J \pm 1.96\,\widehat\tau_J$ for γ_J. The final confidence interval for p_J is then found by transforming to the logistic scale.

(d) The point is to compare regression estimates for the coefficients, one by one, for the two populations, say A and B, using approximate normality. If there is no difference between $\beta_{j,A}$ and $\beta_{j,B}$, then $D_j = \widehat\beta_{j,A} - \widehat\beta_{j,B}$ is approximately normal, with zero mean, and variance $\sigma_j^2 = \sigma_{j,A}^2 + \sigma_{j,B}^2$, with these numbers from the standard errors in the table. For β_1, for example, we have $D_1 = 0.141 - 0.117 = 0.024$, and $\sigma_1 = (0.029^2 + 0.026^2)^{1/2} = 0.0389$, with ratio $D_1/\sigma_1 = 0.616$, too small in absolute value to be significant. The same goes for the other parameter differences, one by one.

Exercise 3: getting children (and children (and children))

(a) Confidence intervals for μ_1, μ_2, μ_3 are found in the traditional t-distribution fashion, with $\bar x \pm t_0\widehat\sigma/\sqrt{n}$, using $t_0 = 1.972$ for the 0.975 quantile of the t_{n-1}. These 95 percent intervals become $[2.949, 3.105]$, $[3.026, 3.180]$, $[3.060, 3.228]$, and are pretty close, with overlap.

(b) The traditional methods for comparing group means, such as the t test and F test, assume independence, which is not realistic here; siblings are correlated via their mothers (and presumably their fathers). The differenes D_i are however safe, in this regard, as these are independent. Testing $\mu_1 = \mu_3$ is the same as testing whether these D_i have mean zero. One finds $t = \bar D/(\widehat\sigma/\sqrt{n}) = 0.117/0.0477 = 2.452$, which is big enough on the standard normal scale to reject the mean zero hypothesis; yes & behold, the 3rd child is significantly bigger than the 1st child, on average.

(c) In such a model, β is the average increase in birthweight, from one child to the next. Again, we cannot analyse data in the usual regression setup, since these assume independence.

(d) The variance of $Y_{i,j}$ is the sum of variances, i.e. $\tau^2 + \sigma^2$. The covariance between weights of two siblings, in this model, is $\mathrm{cov}(M_i + \varepsilon_{i,1}, M_i + \varepsilon_{i,2}) = \mathrm{cov}(M_i, M_i) = \tau^2$.

(e) We test $\beta = 0$ using the ratio statistic $R = \widehat\beta/\widehat\kappa = 0.058/0.023 = 2.522$. Under the null hypothesis this R should be the outcome of an approximate standard normal distribution. But $R = 2.522$ is clearly too big for a standard normal, so $\beta = 0$ is rejected, in favour of $\beta > 0$; birth order matters. The estimated correlation between siblings is $\widehat\rho = \widehat\tau^2/(\widehat\sigma^2 + \widehat\tau^2) = 0.369$.

Exercise 4: bad-tempered and good-tempered husbands and wives

(a) Note that the setup, and questions, are quite similar to those met in one of the Oblig II exercises (with a certain 5×3 table, testing for independence of two factors). From $p_{i,j} = P(X = i, Y = j)$ we have $a_i = p_{i,0} + p_{i,1} = P(X = i)$, for the husbands, and $b_j = p_{0,j} + p_{1,j} = P(Y = j)$, for the wives.

(b) Independence means $P(X = i, Y = j) = P(X = i)P(Y = j)$, i.e. $p_{i,j} = a_i b_j$, for the four possibilities.

(c) One finds

$$\begin{aligned}
\widehat{a}_0 &= (N_{0,0} + N_{0,1})/n = 51/111 = 0.459,\\
\widehat{a}_1 &= (N_{1,0} + N_{1,1})/n = 60/111 = 0.541,\\
\widehat{b}_0 &= (N_{0,0} + N_{1,0})/n = 58/111 = 0.523,\\
\widehat{b}_1 &= (N_{0,1} + N_{1,1})/n = 53/111 = 0.477.
\end{aligned}$$

From these we then get

$$\begin{aligned}
\widehat{E}_{0,0} &= n\widehat{a}_0\widehat{b}_0 = 111(51/111)(58/111) = 26.65,\\
\widehat{E}_{0,1} &= n\widehat{a}_0\widehat{b}_1 = 51 \cdot 53/111 = 24.35,\\
\widehat{E}_{1,0} &= n\widehat{a}_1\widehat{b}_0 = 60 \cdot 58/111 = 31.35,\\
\widehat{E}_{1,1} &= n\widehat{a}_1\widehat{b}_1 = 60 \cdot 53/111 = 28.65.
\end{aligned}$$

(d) The classic test for independence between factors X and Y is the Pearson chi-squared,

$$K = \sum_{i,j} \frac{(N_{i,j} - E_{i,j})^2}{E_{i,j}}.$$

The four terms are all rather small in size, since the $E_{i,j}$ are not far from the observed $N_{i,j}$. Computing gives $K = 1.019$, a number which then should be compared to its distribution under the null hypothesis of independence, which is the chi-squared distribution with degrees of freedom $(r-1)(s-1) = 1$, in the notation for such tables with r rows and s columns. The 0.95 quantile of that distribution is 3.841, so having observed 1.019 is not at all a surprising value. There is no reason to reject the hypothesis of independence, regarding bad-temperedness and gender.

Feel free to collect more data to check if independence holds up with a bigger sample size. Also, Norwegians anno 2024 might not be quite like British couples of the 1880ies.